Thermodynamic Limits on the Size and Size Distribution of Nucleic Acids Synthesized in Vitro: The Role of Pyrophosphate Hydrolysis†

Leonard Peller

ABSTRACT: The free-energy change of phosphodiester bond formation from nucleoside triphosphates is more favorable than with nucleoside diphosphates as substrates. Base-stacking interactions can make significant contributions to both $\Delta G°'$ values. Pyrophosphate hydrolysis when it accompanies the former reaction dominates all thermodynamic considerations. Three experimental situations are discussed in which high-molecular-weight polynucleotides are synthesized without a strong driving force for covalent bond formation. For one of these, a kinetic scheme is presented which encompasses an early narrow Poisson distribution of chain lengths with ultimate passage to a disperse equilibrium population of chain sizes. Hydrolytic removal of pyrophosphate expands the time scale for this undesirable process by a factor of 10^9, while it enormously elevates the thermodynamic ceiling for the average degrees of polymerization in the other two examples. The electron micrographically revealed broad size population from an early study of partial replication of a T7 DNA template is found to adhere (fortuitously) to a disperse most probable representation. Some possible origins are examined for the branched structures in this product, as well as in a later investigation of replication of this nucleic acid. The achievement of both very high molecular weights and sharply peaked size distributions in polynucleotides synthesized in vitro will require coupling to inorganic pyrophosphatase action as in vivo.

If there has been a central activity in molecular biology, it is probably the study of the synthesis of macromolecules. A widely sought objective of this activity is the delineation of extracellular conditions where biologically competent nucleic acids and proteins can be produced from their monomeric precursors.

The investigative efforts in this general area have followed the course of uncovering protein factors which combined with known enzymatic species and small molecule substrates can reconstitute the intracellular synthetic apparatus. By contrast, the quantitative investigation of the purified enzymes and substrates alone in what must admittedly be simplified systems has received scant attention.

The cell-free synthesis of infectious bacteriophage nucleic acid was demonstrated about a decade ago (Spiegelman et al., 1965; Goulian et al., 1967). These RNA and DNA templates and their single-stranded replicated products had molecular weights of $1-2 \times 10^6$. These species are at the small end of the size spectrum of naturally occurring polynucleotides. This also appears to be the size range of the products of partial replication of much larger DNA templates in *well defined* in vitro decoctions.

We feel that thermodynamic limitations are crucially involved in the customary in vitro synthetic systems. Consequently, this communication begins with a review of the free-energy changes accompanying phosphodiester bond synthesis in single- and double-stranded chains. The question of the thermodynamic restrictions imposed on the average size of the chains is then considered. A mechanistic scheme is presented for primer-initiated single-strand polynucleotide formation and the formal kinetics are developed to illustrate the conversion of a narrow size distribution to a wide one with increasing time. Finally, the broad range of chain lengths of a partially replicated bacteriophage DNA of moderate size is quantitatively examined.

Our emphasis is on the role of the relatively irreversible hydrolysis of the pyrophosphate synthetic by-product in setting a high ceiling on the molecular weights for polynucleotide syntheses and in deferring the time of decay of the initially sharp spectrum of chain sizes.

Thermodynamics of Bond Formation in Nucleic Acid Synthesis

The fundamental reaction for the addition of a nucleotide unit to a polyribonucleotide chain utilizing a ribonucleoside diphosphate (rNDP) as the monomer source and yielding an orthophosphate ion product (P) can be written as

$$I \cdots N_pN_pN + rNDP \rightleftharpoons I \cdots N_pN_pN_pN + P \quad (1)$$

At the 5′ end of this chain is an oligonucleotide primer moiety designated as I, which can only be formed slowly, if at all, by the polyribonucleotide phosphorylase catalyzing this reaction. The facile reversibility of this polymerization has long been recognized (Grunberg-Manago and Ochoa, 1955). A standard free-energy change $\Delta G°_1' \approx 0$ is appropriate for this reaction (Grunberg-Manago, 1963; Peller and Barnett, 1962). The small departures from this isoergonic condition can be accounted for by a negative contribution to the free energy of polymerization arising from single-strand base–base interactions (Peller, 1976a).

The analogous reaction for extension by one nucleotide unit of a growing polydeoxyribonucleotide chain with a deoxyribonucleoside triphosphate (dNTP) as the monomeric source and yielding pyrophosphate (PP) as the immediate by-product is given by

$$I \cdots N_pN_pN + dNTP \rightleftharpoons I \cdots N_pN_pN_pN + PP \quad (2)$$

Again the 5′ terminus is occupied by a primer entity (I) which bears a similar relationship of limited susceptibility to depolymerization as in the previous reaction but now to terminal deoxynucleotidyltransferase which catalyzes the above reaction

† From the Cardiovascular Research Institute, University of California, San Francisco, California 94143. *Received August 25, 1976.* This research was supported by U.S. Public Health Service Program Project Grant HL-06285 and by the Samuel R. Neider Heart Research Fund.

(Krakow et al., 1961; Bollum, 1962). The free-energy change for this addition can be related to that for reaction 1 and the sum and difference, respectively, of the free energies of hydrolysis of ATP to form AMP and PP and ADP to provide AMP and P (Peller, 1966). The best available data for these latter two quantities (Alberty, 1969) suggest a value of $\Delta G^{\circ}{}_2' = -1$ kcal (Peller, 1976b), again ignoring base-stacking interactions. The directly determined equilibrium constant of approximately 100 for deoxyadenosine triphosphate as a substrate (Kato et al., 1967) can be shown to diverge quantitatively from this estimate (Peller, 1976a) because of these interactions as gauged by the marked tendency of deoxyadenosine to associate in aqueous solution (Solie and Schellman, 1968). We emphasize at this point that the actual magnitude of these free-energy changes within rather substantial limits has little effect on the arguments to be developed below.

The formation of double-stranded nucleic acids from deoxyribonucleoside triphosphates can be viewed as the sum of two reactions like the above, followed by an association process involving two i-meric complementary chains

$$\cdots {}_{p}N_{p}N_{p}N\cdots + \cdots N_{p}'N_{p}'N_{p}'\cdots \rightleftharpoons \begin{matrix}\cdots {}_{p}N\ {}_{p}N\ {}_{p}N\cdots \\ | \quad | \quad | \\ \cdots N_{p}'N_{p}'N_{p}'\cdots\end{matrix} \qquad (3a)$$

The effective equilibrium constant for this reaction $\sigma_0 s^{i-1}$ involves a constant for the establishment of the first hydrogen-bonded pair (σ_0) and a base-pairing propagation constant (s) which figures in the coil to helix transformation (Zimm, 1960). As has been noted earlier, $\sigma_0 s^{i-1}$ is a minimum value for this association reaction as partially hydrogen-bonded intermediate states are neglected (Peller, 1966). No mechanism of replication is subsumed in this convenient thermodynamic representation. The free-energy change of the growth step for a *single* strand of the product bihelix is given by

$$\Delta G^{\circ}{}_3' = \Delta G^{\circ}{}_2' - (RT/2)\ln s = -(RT/2)\ln (K_2{}^2 s) \qquad (3b)$$

where contributions due to covalent bond formation (reaction 2) and base–base interactions are separated. A mean value $\Delta G^{\circ}{}_3' \approx -2.5$ kcal has been estimated for this process (Peller, 1976b) with a considerable range due to the variation of bihelical base-stacking interactions with nucleotide sequence (Borer et al., 1974).

The free-energy changes of the chain-growth steps for reactions 1–3 pale significantly compared to that for the highly exergonic hydrolysis of the pyrophosphate catalyzed by inorganic pyrophosphatase

$$PP + H_2O \rightleftharpoons 2P \qquad (4)$$

A recent compilation of data (Alberty, 1969) gives for this reaction $\Delta G^{\circ}{}_4' = -7.5$ kcal.

Enzyme-catalyzed polymerizations in vitro are subject to thermodynamic limitations with respect to the number of bonds which can be formed. Such reactions occurring in *closed* systems with no admission of either substrate or removal of product can simply not synthesize more bonds than the final equilibrium state permits. The overall number average degree of polymerization must then at all times be less than this limiting equilibrium value. This constraint has gone unrecognized so that the question of the distribution of a restricted number of bonds over a variable number of chains has not commanded serious examination.

Equilibrium Size Limits for the In Vitro Production of Polynucleotides

It is more than a little surprising that readily reversible reactions, such as those cited above, can lead to high-molecular-weight ($\sim 10^6$) polynucleotide chains. The early availability of such species in these systems has served as a distraction from any scrutiny of the conditions for their existence. The prevalent wisdom has been that macromolecules are a reasonable consequence of the action of polymerizing enzymes on "high-energy" substrates.

In summarizing and extending some earlier results, we point out that this picture is far from satisfactory. Specifically, three realizable situations are definable where large polymers can arise if transiently without the benefit of very exergonic synthetic steps. The relevant thermodynamic information permits conclusions as to which situations lead to essentially *stable* high-molecular-weight products. Examples of all can be found with in vitro polynucleotide syntheses utilizing purified enzymes.

(*i*) *Polynucleotide Synthesis on a Preexistent Primer.* Let the synthesis of polynucleotides occur *exclusively* by addition of nucleotide units to a fixed number of preexistent primer centers, as in reactions 1 and 2. Then, the equilibrium number average degree of polymerization ($\langle i \rangle$) for chains so generated is given for the first example by

$$\langle i \rangle = 1 + (K_1/K_1 + p)\{[\text{rNDP}]_0/[\text{I}]_0\} \approx (K_1/K_1 + 1)\{[\text{rNDP}]_0/[\text{I}]_0\} \qquad (5a)$$

where $[\text{rNDP}]_0$ and $[\text{I}]_0$ are the total concentrations of "activated" monomer ribonucleoside diphosphate and primer oligonucleotide, respectively (Peller and Barnett, 1962). The equilibrium probability (p) of phosphodiester bond formation is related to the other polymerization parameters by

$$p \equiv K_1([\text{rNDP}]_e/[\text{P}]_e) = K_1([\text{rNDP}]_e)/\{([\text{rNDP}]_0 - [\text{rNDP}]_e)\} \approx \{([\text{rNDP}]_0)/([\text{rNDP}]_0 + [\text{I}]_0)\}\{1 - ([\text{rNDP}]_0[\text{I}]_0)/K_1([\text{rNDP}]_0 + [\text{I}]_0)^2\} \qquad (5b)$$

where the subscript e refers to the concentrations at equilibrium. Now the total number average degree of polymerization including all the unreacted monomer is given by

$$\langle i \rangle_T \approx (K_1 + 1)[\text{rNDP}]_0/\{(K_1 + 1)[\text{I}]_0 + [\text{rNDP}]_0\} \qquad (5c)$$

The condition for achieving high degrees of polymerization of the chains growing out from the 3′ OH of the primer oligonucleotide is readily seen from eq 5a. The first factor, $K_1/(K_1 + 1)$, in this expression is the order of 1/2 for $\Delta G^{\circ}{}_1' \approx 0$. Moreover, this quantity, which can be approximately identified with the equilibrium extent of incorporation of nucleotide into polymer, can never exceed one irrespective of the magnitude of K_1. A large value of $\langle i \rangle$ requires a high ratio of diphosphate monomer source to primer, which is the second factor in eq 5a.

If the condition that $[\text{rNDP}]_0 \gg [\text{I}]_0$ applies, then from eq 5c $\langle i \rangle_T \approx K_1 + 1 \approx 2$. This is, of course, simply another way of stating that a significant amount of monomer source remains unreacted.

For reaction 2 with a more negative free-energy change, substantially the same situation applies even if a value of $K_2 \approx 100$ is chosen. Thus, values of $[\text{dNTP}]_0/[\text{I}]_0 \approx 10^3 - 10^4$ must be presumed for number average molecular weights of 10^5 to 10^6 daltons. Furthermore, as this ratio represents the upper limit to the number average degree of polymerization of the primer-initiated chains, $\langle i \rangle$, no increase in this quantity can be attained by coupling reaction 2 to the hydrolysis of the PP product represented by reaction 4. Driving the polymerization to complete incorporation of triphosphates, however, does lead to $\langle i \rangle_T = \langle i \rangle = [\text{dNTP}]_0/[\text{I}]_0$. Some further im-

portant consequences of this coupling process will be developed below.

(*ii*) *De Novo Synthesis of Single-Strand Polynucleotides.* The conversion of prescribed nucleotide sequences of DNA into RNA transcripts belongs in this category. The RNA chains so synthesized are generally commenced by a purine nucleoside triphosphate (Bremer et al., 1965) but appear not to require a preformed oligonucleotide primer (Geiduschek and Haselkorn, 1969). Moreover, the transcripts are only transiently associated with the imprinting DNA. From a thermodynamic standpoint this can be viewed as equivalent to de novo single-strand polynucleotide formation.

Unlike case i, there is no easy rationale for the separation of monomeric triphosphate species from oligomeric and polymeric chains terminated by a 5′ triphosphate moiety. However, if we exclude the monomeric ribonucleoside triphosphates from the summation of species in analogy with the above, then the equilibrium number average degree of polymerization of all chains containing one or more phosphodiester bonds is given by

$$\langle i \rangle = 2 + K_2^{1/2} \tag{6a}$$

By utilizing the equilibrium constant for reaction 2, we here ignore any difference between ribonucleoside and deoxyribonucleoside triphosphates. The probability for chain growth can be written as

$$p = K_2^{1/2}/(1 + K_2^{1/2}) \tag{6b}$$

(Peller and Barnett, 1962). The number average degree of polymerization, including the monomeric reactant, is in turn

$$\langle i \rangle_T = 1 + K_2^{1/2} \tag{6c}$$

We have recapitulated these results of simple mass-action considerations (Peller, 1975) primarily to contrast them with the oligonucleotide-primed single-strand polynucleotide formation of case i and the do novo synthesis of the double-strand species below. Firstly, we note that the equilibrium extent of polymerization of any selected population of chains above some minimal size will not differ greatly from either $\langle i \rangle$ or $\langle i \rangle_T$. Secondly, for even impossibly large values of K_2, e.g. 10^4, the limiting number average degrees of polymerization calculated from eq 6a and 6c are really quite small. Lastly, if a reasonable upper limit of $K_2 \approx 100$ is chosen, the final extent of incorporation of ribonucleoside triphosphates into polymer is from eq 6b only 90%, i.e., measurably incomplete.

An earlier argument (Peller, 1975) has shown that for the equilibrium reached in the presence of inorganic pyrophosphatase, K_2 in eq 6a–c is multiplied approximately by the factor $K_4/4[\mathrm{rNTP}]_0$ appropriate to the one-sided equilibrium of reaction 4. With customary concentrations of precursor triphosphate (10^{-4} M) and the large negative value of $\Delta G°_4{}'$ quoted above, this factor is the order of 10^9. There is a corresponding very large elevation of the ceiling of the thermodynamic average degrees of polymerization by a factor of 10^4 to 10^5.

In vitro transcription of bacteriophage T7 DNA has been reported to yield chains of mRNA size, i.e., 10^5–10^6 daltons (Dunn and Studier, 1973), *without* accompanying hydrolysis of the pyrophosphate by-product. We emphasize that this must be a purely kinetic phenomenon. The presence of very long chains at early stages in the synthesis in all likelihood derives from a chain-initiation rate, which is markedly slower than that of chain longation. A factor of 10 to 20 in relative rates for these two quite different steps seems to be suggested from some presently available data (Chamberlin, 1974). Case i discussed above represents a limit where de novo chain initiation is negligibly slow and macromolecular synthesis must come entirely from preformed initiator elements.

There is a tendency for the rate of net RNA syntheses in vitro to diminish with increasing time not merely from the depletion of substrates (Bremer, 1967; Hyman and Davidson, 1970) but, we believe, due to the onset of pyrophosphorolysis. Under these conditions, the high-molecular-weight early products are labile and subject to a redistribution of their nucleotide units to drastically smaller sized species arising from the steady advent of newly initiated chains.

(*iii*) *De Novo Synthesis of Double-Strand Polynucleotides.* Examples of the de novo production of duplex polynucleotides, which contain a single nucleotide in one strand and its complementary nucleotide in the associated strand, are provided for DNA synthesis by dG:dC (Radding et al., 1962) and RNA synthesis by rA:rU (Smith et al., 1967; Krakow et al., 1968). The equilibrium state for this process has been considered previously (Peller, 1966). We summarize some of these results in the format used for cases i and ii.

The number average degree of polymerization of bihelical product was found to be given by

$$\langle i \rangle \approx K_2(s/\sigma_0[\mathrm{NTP}]_0)^{1/2}[1 + (K_2 s^{1/2})^{-1}]$$
$$\approx K_2(s/\sigma_0[\mathrm{NTP}]_0)^{1/2} = K_3(\sigma_0[\mathrm{NTP}]_0)^{-1/2} \tag{7a}$$

with the two reactant nucleoside triphosphates having the same initial concentration. The probability of *single*-strand elongation can be expressed as

$$p \equiv K_2([\mathrm{NTP}]_e/[\mathrm{PP}]_e) \approx s^{-1/2} \tag{7b}$$

The corresponding probability for growth of the bihelix is $p^2 s$ and the final approximation in eq 7b applies when the associated state is the more stable ($s > 1$) and high molecular weights of this product occur ($p^2 s \approx 1$). The total number average degree of polymerization, including monomeric triphosphates as well as all single-strand species and the bihelical products counted in eq 7a, is

$$\langle i \rangle_T \approx 1 + (K_2 s^{1/2})/2 \approx (K_2 s^{1/2})/2 = K_3/2 \tag{7c}$$

It is convenient to discuss the driving force for the polymerization in terms of the product of three terms in eq 7a. The first factor, K_2, representing covalent bond formation *may* be as large as 100, while the third factor is of the order of unity. The second factor, $(s/\sigma_0[\mathrm{NTP}]_0)^{1/2}$, reflects the development of the secondary structure in the product as the ratio of equilibrium constants for propagation and initiation of the bihelix. This is a cooperative feature of the synthesis involving the energetics of hydrogen bond formation and base stacking.

With $K_3 \approx 100$ derived from $\Delta G°_3{}' = -2.5$ kcal quoted above, the final form for $\langle i \rangle$ in eq 7a demonstrates that the association equilibrium constant, σ_0, must be *small* to provide a high thermodynamic limit for the extent of polymerization. With $\sigma_0 \approx 10^{-2}$ obtained from the acid association of oligoriboadenylic acid (Applequist and Damle, 1965) and $[\mathrm{NTP}]_0 = 10^{-4}$ M, we find $\langle i \rangle \approx 10^5$ or a molecular weight of approximately 60×10^6, the same order of magnitude as an earlier calculation (Peller, 1966). From more recent determinations of the association constant for oligomers of complementary ribonucleotides, values of $\sigma_0 \approx 10^{-4}$ have been reported (Borer et al., 1974). With this estimate for the initiation parameter, the limiting value of $\langle i \rangle$ would be an order of magnitude higher, i.e., 10^6. The absence of Mg^{2+} in these

studies may lead to a misleadingly low estimate of σ_0 (Peller, 1976b) and a corresponding overestimate of the limiting molecular weights. Even the lower of these two calculated limiting degrees of polymerization is substantially larger than the in vitro partially replicated DNA cited above. However, both calculations are well below the size of bacterial DNAs let alone the stupendous polynucleotide chains composing eucaryotic chromosomes (Kavenoff and Zimm, 1973).

The cooperative aspect of the synthesis involving secondary forces clearly predominates over the driving force assigned to covalent bond formation. In a sense, the factor $(s/\sigma_0[\mathrm{NTP}]_0)^{1/2}$ should be compared with $[\mathrm{rNDP}]_0/[\mathrm{I}]_0$ in eq 5a, which plays a similar causative role in the origin of a high-molecular-weight product in case i. A low concentration of primer is roughly equivalent to a small value of σ_0 making initiation of a double helix difficult. As an additional parallel to case i, the total number average degree of polymerization given by eq 7c depends essentially on the thermodynamics of phosphodiester bond formation and can be quite small even when $\langle i \rangle$ is very large.

The principal difference between cases ii and iii is the cooperative element which is missing in single-strand synthesis. From the complete formulation of case iii (Peller, 1966), substantially the results of case ii are recovered when single strands are stable ($s < 1$). The quantitative importance of a nucleation step in polymerization was first recognized for protein-association reactions (where only secondary forces are involved) like the formation of F-actin double helices (Oosawa and Kasai, 1962).

With the coupling of pyrophosphate hydrolysis to the syntheses of double-helical polynucleotides, K_3 in eq 7a is multiplied by the immense factor $K_4/4[\mathrm{NTP}]_0$. This yields an astronomical ceiling for the equilibrium degree of polymerization of from 10^{13} to 10^{14}! This hydrolytic process has sometimes been apologetically described as "wasteful", except for the benefit that under intracellular conditions the synthetic product will enjoy a freedom from fluctuations in the level of pyrophosphate concentration (Kornberg, 1962, 1974a). Below we will argue for both a limiting average degree of polymerization vastly higher than the template species as well as a complete conversion of nucleoside triphosphate to polynucleotide in vitro with the total abolition of all vestiges of a depolymerization reaction.

In concluding this summary of the thermodynamic limits in nucleic acid synthesis, certain generalizations can be proffered. Of the three cases considered, the high-molecular-weight product in i will be stable so long as there is no further introduction of initiator oligonucleotide, while the polynucleotides in ii are clearly kinetically unstable. The macromolecules synthesized in case iii represent a kinetically stable state but with their existence crucially predicated on a small probability of nucleation of a double helix. We must note here that there is a widely held disbelief in the capacity of enzymes catalyzing DNA synthesis to initiate new chains (Schekman et al., 1974). This has led to a mechanistic proposal for "de novo" synthesis of an alternating dAT copolymer which is dependent on the continuous copying of an adventitiously present oligomer of this type bound to the enzyme albeit in infinitesimal amount (Kornberg et al., 1964). This explanation for the origin of a self-complementary polynucleotide is not readily extended to the synthesis of complementary homopolymers, such as dG:dC.

The consideration of these equilibria posits that the reverse reactions cannot be neglected. The depolymerization or phosphorolysis reaction in i received early examination and the pyrophosphorolysis of DNA iii has been more recently demonstrated (Beyersmann and Schramm, 1968; Deutscher and Kornberg, 1969). Reversibility of RNA synthesis (ii) appears to have passed unrecognized.

There is irony in the pursuit of conditions for the facile initiation of chains in nucleic acid synthesis (Schekman et al., 1974). The one common feature of all three cases is that the syntheses of macromolecules in the absence of a strongly exergonic covalent-bond-forming reaction critically depend on kinetic or thermodynamic inhibition in the initial steps.

Size Distributions in Primer-Initiated Polynucleotide Synthesis

The foregoing discussion has dealt with the thermodynamics of nucleic acid synthesis and the limits imposed on the size of the in vitro product. In examining the thermodynamic basis or a lack of same for these polymerizations, we have sidestepped the issue of the distribution of chain sizes. The equilibrium limit requires that this size range be quite broad and describable as a geometric or most probable population (Flory, 1953).

Thermodynamic arguments possess the obvious advantage of being independent of mechanistic considerations. However, with a given set of reactants, rather different products can be obtained depending on the presumed kinetic admissibility of the final equilibrium states, viz., cases ii and iii above. In order to investigate the synthetic process in further detail, a mechanism for the polymerization must be proposed and a kinetic formulation developed. Case i above with pyrophosphate as by-product is chosen for examination because it is, by far, the most tractable and has clear precedents in the realm of synthetic high polymers.

A rather simple scheme for the conversion of a chain length of $i-1$ to one of i nucleotides can be written as comprising five sequential steps:

$$\mathrm{E} + \mathrm{NTP} \rightleftharpoons \mathrm{ENTP} \qquad \text{(I)}$$

$$\mathrm{ENTP} + \mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i-1} \rightleftharpoons \mathrm{ENTP} \cdot \mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i-1} \qquad \text{(II)}$$

$$\mathrm{ENTP} \cdot \mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i-1} \rightleftharpoons \mathrm{EPP} \cdot \mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i} \qquad \text{(III)}$$

$$\mathrm{EPP} \cdot \mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i} \rightleftharpoons \mathrm{EPP} + \mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i} \qquad \text{(IV)}$$

$$\mathrm{EPP} \rightleftharpoons \mathrm{PP} + \mathrm{E} \qquad \text{(V)}$$

The sum of steps I–V, of course, yields reaction 2.

Rate expressions can be obtained for this simple compulsory pathway by either a schematic enumeration of terms (King and Altman, 1956) or a direct recursive solution of the algebraic equations governing the steady state in the various enzyme–substrate complexes (Bloomfield et al., 1962). The resultant rate law for the change in concentration of i-mer is composed of four contributions, as a primer extended by i nucleotide units can be formed by the addition of a nucleotide to its predecessor of $i-1$ units and can, in turn, disappear by pyrophosphorolysis to regenerate this species, as indicated in the above sequence while a similar series of steps connects the i-mer to the $i+1$-mer. A concise form for this is

$$\frac{\mathrm{d}[\mathrm{I}(_{\mathrm{p}}\mathrm{N})_i]}{\mathrm{d}t} = v_{\mathrm{f}}[\mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i-1}] - (v_{\mathrm{f}} + v_{\mathrm{b}})[\mathrm{I}(_{\mathrm{p}}\mathrm{N})_i] + v_{\mathrm{b}}[\mathrm{I}(_{\mathrm{p}}\mathrm{N})_{i+1}] \qquad \text{(8a)}$$

for $i \geq 2$ where

$$v_f \equiv \frac{k_{II}k_{III}k_{IV}[E]_0}{k_{-II}k_{IV} + k_{III}k_{IV} + k_{-II}k_{-III}} \times \left\{\frac{[NTP]/K_I}{1 + [NTP]/K_I + [PP]/K_V}\right\} \quad (8b)$$

$$v_b \equiv \frac{k_{-II}k_{-III}k_{-IV}[E]_0}{k_{-II}k_{IV} + k_{III}k_{IV} + k_{-II}k_{-III}} \times \left\{\frac{[PP]/K_V}{1 + [NTP]/K_I + [PP]/K_V}\right\} \quad (8c)$$

The k's with positive subscripts refer to the forward or chain-elongation direction of reaction, while those with negative subscripts apply to the reverse direction or pyrophosphorolysis. K_I and K_V are simply dissociation constants for NTP and PP complexes, respectively, with the enzyme of total concentration $[E]_0$.

The primer oligonucleotide is assumed to be of sufficient length so that the same kinetic parameters govern its reaction as that of its polymeric products, which are independent of i. With this stipulation, the following two equations complete the set (eq 8a)

$$\frac{d[I]}{dt} = -v_f[I] + v_b[I(_pN)] \quad (9a)$$

$$\frac{d[I_pN]}{dt} = v_f[I] - (v_f + v_b)[I(_pN)] + v_b[I(_pN)_2] \quad (9b)$$

These steady-state expressions have been derived with the explicit requirement that $[I]_0 \ll [E]_0 \ll [NTP]_0, [PP]$. The first inequality means that the enzyme can be accounted for largely in three forms—free, complexed with NTP, or complexed with PP. Only the fraction in the first complex, the term in braces in eq 8b, can catalyze the forward reaction, while the term in braces in eq 8c represents the fraction capable of catalyzing pyrophosphorolysis. The premultiplying factors of these two terms are the two maximum velocities for nucleotide addition (V_f) and pyrophosphorolysis (V_b), respectively. The overall rate laws for these polymeric species are clearly of the canonical Michaelis–Menten form. This same form extends to the expression for nucleoside triphosphate consumption, which has been customarily utilized for template-directed RNA synthesis (Bremer, 1967; Hyman and Davidson, 1970).

In the early stages of this process where the reverse reaction can be ignored, i.e., $[PP] \approx 0$ and $v_b \approx 0$, the kinetic representation by eq 8a, 9a, and 9b reduces to that for the alcohol initiated polymerization of ethylene oxide treated many years ago (Flory, 1940; Gold, 1958). The number fraction of species was shown to be of the Poisson type with the same distribution applying approximately to the weight fraction. The solution of the set of differential equations for $[I(_pN)_i]$ with $v_b = 0$ can be expressed in terms of a Poisson weight fraction $W_p(i)$

$$W_p(i) \equiv \{i[I(_pN)_i]\} / \left\{\sum_{j=0}^{\infty} j[I(_pN)_j]\right\} \approx \langle i\rangle^i(\exp - \langle i\rangle)/i! \quad (10a)$$

where

$$\langle i\rangle \approx \int_0^t v_f dt = \{[NTP]_0 - [NTP]\}/[I]_0 \quad (10b)$$

The particular features of the system under discussion are embodied exclusively in eq 10b.

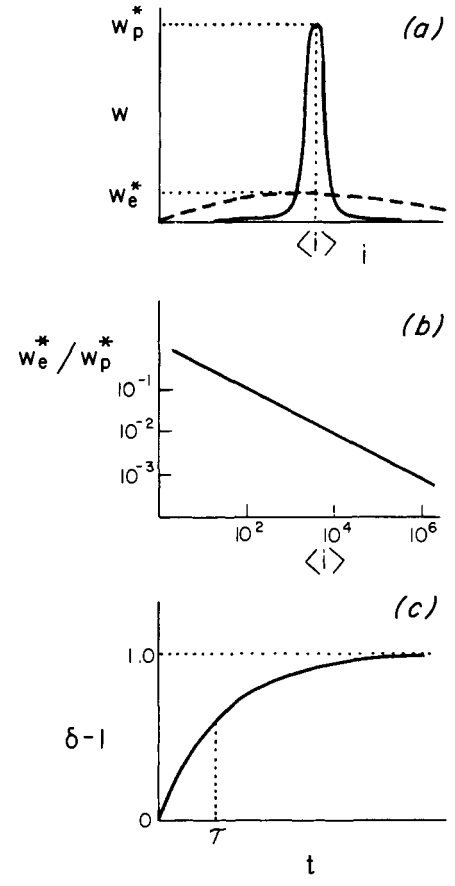

FIGURE 1: Some aspects of the two common limiting size distributions for linear polymers. (a) Schematic depiction of the weight fractions for the most probable or equilibrium (W_e), dashed curve, and Poisson (W_p), solid curve, distributions vs. the degree of polymerization (i). (b) Double-logarithmic plot of the ratio of the two maxima in a W_e^*/W_p^* vs. the number average degree of polymerization ($\langle i\rangle$). (c) Representation of the exponential time course for the conversion of the polydispersity index (δ) from the Poisson state ($\delta = 1$) to the final equilibrium plateau ($\delta = 2$).

When all the derivatives are set equal to zero, the equilibrium result follows with

$$[I(_pN)_i]_e/[I(_pN)_{i-1}]_e = v_f/v_b = K_2[NTP]_e/[PP]_e \equiv p \quad (11)$$

K_2 is the equilibrium constant for reaction 2 and p is given by the appropriate analog of eq 5b. The expressions comporting with eq 10a and 10b but for the parameters of the most probable distribution arising in this limiting case are

$$W_e(i) = (i + 1)p^i(1 - p)^2 \approx (i/\langle i\rangle^2)(\exp - i/\langle i\rangle) \quad (12a)$$

$$\langle i\rangle = p/(1 - p) \approx (K_2/K_2 + 1)\{[NTP]_0/[I]_0\} \quad (12b)$$

The final approximate form for the equilibrium weight fraction (W_e) applies when $p \rightarrow 1$ and represents a continuous variation of i (Morawetz, 1965). Again the particularities of this system reside entirely in eq 12b for the number average degree of extension of the polynucleotide chain.

These two distributions arising in the extreme situations presented above have quite different quantitative properties. Figure 1a exhibits these differences qualitatively in the fashion of texts (Flory, 1953; Tanford, 1961). For sufficiently high degrees of polymerization ($\langle i\rangle > 100$), both weight fraction distributions have maxima at $i^* = \langle i\rangle$ but that corresponding to the most probable or equilibrium distribution is much the broader. The equilibrium most probable distribution broadens with increasing $\langle i\rangle$ as the latter appears reciprocally in the

exponent of eq 12a, while the Poisson narrows further as is evident from eq 10a. This divergent behavior is depicted in Figure 1b, where the logarithm of the ratio of the two maxima is plotted against the logarithm of $\langle i \rangle$ (for sufficiently high degrees of polymerization $W_e^*/W_p^* \approx \langle i \rangle^{-1/2}$).

The above results have a long history in the area of synthetic high polymers. However, it is only in more recent years that the kinetics of the passage of a system from an initial sharply peaked population to a final highly disperse range of sizes has been analyzed (Brown and Szwarc, 1958; Miyake and Stockmayer, 1965; Oosawa, 1970; Nanda and Jain, 1970). Only approximate results are available for this process, but the conclusions of one of these studies (Miyake and Stockmayer, 1965) can be simply summarized in the context of this system. First, the time scale for the attainment of the equilibrium ratio of nucleoside triphosphate to pyrophosphate product ($[NTP]_e/[PP]_e \approx K_2^{-1}$ from eq 11) can be orders of magnitude shorter than the time of appreciable reshuffling of monomer units between chains to yield the ultimate broad equilibrium size range ($[I(pN)_{i-1}]_e \approx [I(pN)_i]_e$). For the latter, the kinetics describe an approximately exponential approach of the polydispersity index $\delta \equiv \langle i^2 \rangle / \langle i \rangle^2$ from an initial value of unity to a final value of 2, schematically shown in Figure 1c as

$$\delta - 1 \approx 1 - (\exp - t/\tau) \tag{13}$$

where the time constant, τ, is given by

$$\tau = \langle i \rangle_e^2 / 2(v_b)_e \tag{14a}$$

(Miyake and Stockmayer, 1965). The magnitude of $\langle i \rangle_e$ is determined effectively by the ratio of triphosphate to initiator species, as indicated in the previous section. $(v_b)_e$ is the reverse rate "constant" evaluated with the equilibrium concentrations, i.e., $[PP]_e$ and $[NTP]_e$. Consequently,

$$\tau \approx \{(K_2/K_2 + 1)^2([NTP]_0/[I]_0)^2(K_V)(1 + [NTP]_e/K_I + [PP]_e/K_V)\}/2V_b[PP]_e \approx \{([NTP]_0/[I]_0)^2\}/2V_b \tag{14b}$$

where the condition that $[PP]_e \approx [NTP]_0 \approx K_V$ has been invoked as the final approximation. Even an incomplete conversion of substrate to polynucleotide will not invalidate this approximation, provided that $[NTP]_0$ is sufficiently large.

For the polymerization process coupled to the immediate or nonrate-limiting hydrolysis of pyrophosphate by inorganic pyrophosphatase, then $[NTP]_e$ *is immeasurably different from zero,* while $[PP]_e \approx 4[NTP]_0^2/K_4 \approx 10^{-12}$ M for $[NTP]_0 \approx 10^{-4}$ M—an infinitesimal quantity compared to the customary Michaelis constants for small substrates bound to enzymes. Consequently, the redistribution time constant under these conditions, τ', is given by

$$\tau' \approx ([NTP]_0/[I]_0)^2 K_4 K_V/8V_b[NTP]_0^2 \tag{15a}$$

The ratio of these two parameters is then

$$\tau'/\tau \approx K_4 K_V/4[NTP]_0^2 \approx K_4/4[NTP]_0 \tag{15b}$$

under initial conditions of partial saturation of the enzyme by substrate, viz., $[NTP]_0 \approx K_V$.

The above argument justifies on plausible kinetic grounds the introduction of the ubiquitous large factor arising from the relatively irreversible hydrolysis of pyrophosphate. *If τ were the order of an hour, τ' would from eq 15b be the order of 10^4 years.* The discussion of case i above noted that the limiting extent of polymerization could not be much influenced by a highly favorable equilibrium, but certainly the likelihood of delaying the onset of an undesirable size dispersal is enormously enhanced.

The proposal has been made that the primer-initiated synthesis of polydeoxyribonucleotides *without* the concomitant hydrolysis of the pyrophosphate fulfills the conditions originally set forth for the procurement of a Poisson distribution (Bollum, 1974). However, the original analysis (Flory, 1940) specifically precluded the intervention of a depolymerization reaction. The reaction catalyzed by terminal deoxyribonucleotidyltransferase is measurably reversible. Equilibrium constants for nucleoside triphosphate substrates containing bases with a small tendency to interact with chain neighbors will doubtless be less than that reported for dATP cited in this proposal.

The removal of pyrophosphate *has* been recommended for this system to diminish the effect of product inhibition on the reaction rate (Bollum, 1974). Product inhibition, as customarily defined in the kinetics of enzyme reactions yielding two or more products, arises on the formation of an enzyme complex with one of these species alone when the others are absent (Alberty, 1958; Bloomfield et al., 1962). Thus, the enzyme can be largely converted to a form ineffective in catalyzing the forward reaction but still incapable of catalyzing the reverse reaction because of the lack of the other product molecules. In simple algebraic terms for the present kinetic scheme, such an effect diminishes the polymerization rate by increasing the denominator of v_f, while the effect of the pyrophosphate in augmenting the rate of depolymerization is by increasing the numerator of v_b. Product inhibition in its customary guise is difficult to assess in enzyme-catalyzed polymerizations because the second product, the polymer chain, functions as both a reactant and a product. However, if the pyrophosphate acted exclusively in this inhibitory way, it would not alter the prospects for obtaining a Poisson distribution. Only the magnitude of v_f would be modified and, hence, the change of $\langle i \rangle$ with time as given by eq 10b.

This oversimplified enzyme kinetic scheme was adopted because it has permitted a straightforward paraphrase of some of the results of the statistics of synthetic polymer formation. In particular, the two extreme distributions for linear polymers are readily embraced. In this connection, the observation is of interest that a limited addition of nucleotides to an exogenous oligomeric primer by terminal deoxyribonucleotidyltransferase to yield an $\langle i \rangle = 11$ follows a Poisson distribution (Hayes et al., 1967). Equation 14b would argue that such a state with a small average size would persist with difficulty unless the reverse reaction were really quite negligible.

There is an additional more insidious problem with many of these synthetic reactions. This concerns the shadowy nature of the initiator species presumed to engender the polymerization. Most often this primer is not an oligomer added to the system but rather an endogenous contaminant of the enzyme. In contradistinction to the alcohol initiator of the synthesis of polyethylene oxide or the amine initiator of the polymerization of N-carboxy anhydrides (Sela and Katchalski, 1958), this initiator does not differ molecularly from the polymeric product, except in chain length. There may then be good reason to anticipate that the enzyme has the catalytic capacity to form such a primer species albeit rather slowly compared to its subsequent growth. Such appears to be the case with polynucleotide phosphorylase (Godefroy-Colburn and Grunberg-Manago, 1972). If the reverse reaction were to operate measurably, the newly formed oligomers could grow at the expense of the monomeric units of the earlier and larger synthesized chains. This situation would be disastrous for the maintenance

of both a high-molecular-weight product and a narrow size distribution. For reactions with pyrophosphate as a product, the avoidance of this condition provides another argument for its hydrolytic removal.

Size Distributions in Partially Replicated DNA

The two distributions in the preceding development provide a reference frame for the examination of the population of chain lengths arising in the in vitro replication of double-strand nucleic acids. Information on the length and character of the product chains from replication studies of DNA from T7 bacteriophage has been available for some years initially with the enzyme presently designated DNA polymerase I from *Escherichia coli* (Inman et al., 1965). None of the ancillary factors now believed to be necessary for DNA replication in vivo were known when these studies were made and, hence, are not wittingly present in the system.

From a T7 DNA template which by electron microscopic examination was overwhelmingly a simple linear bihelix, replicated products were obtained, which under the same scrutiny manifested extensive branching. The degree of ramification was reported to have increased with the extent of DNA synthesis. Such structures were also observed in the de novo synthesis of DNA-like polynucleotides (Kornberg, 1965).

Because ramiform products have been widely encountered with in vitro replication studies, consideration must be given to possible mechanisms for their generation. Figure 2 depicts three modes of origin of branched chains in the nascent strands. None of these alternatives require the intercession of an RNA primer, as this phenomenon was first observed under conditions where neither the proper substrates for this initiation mechanism nor, more importantly, the enzyme RNA polymerase was present. The labeled and separated newly replicated strands also exhibited a similar ramified pattern. Consequently, we have attempted to represent the nature of the product after strand dissociation in alkali, separation of the product of different density than the template, and renaturation of these newly synthesized polynucleotide chains.

Figure 2a presumes a chain initiation from an adventitious oligonucleotide DNA primer (short heavy line) attached initially to one of the strands of the template. Repeated replication on the newly generated chain results in frequent involutions or "hairpin loops" in the product. The separated material (lower entry in Figure 2a) is branched but is a *continuous* single strand with bihelical base-paired regions.

The sketch in Figure 2b represents a recent proposal for the origin of ramified bihelical structures (Kornberg, 1974b) issuing from a template in which an endonucleolytic scission is initially present. The lower entry depicts the separated newly synthesized material. Further replication would presumably yield a more highly branched product. This mode of generation of branched *separated* strands *requires* the further intervention of an endonucleolytic or chain-cleavage action to yield the final product. In this postulated mechanism, the second scission occurs near the site of the first, presumably, for the newly synthesized chains to be readily separated by density difference from the template.

The dissociated products in Figure 2a,b are both single stranded containing self-complementary regions interrupted by short nonhelical turns. The ease of renaturation of the product which has been reported is a consequence of its single-strand character. In neither case is it necessary to assume that the enzyme system has the capacity to initiate new chains.

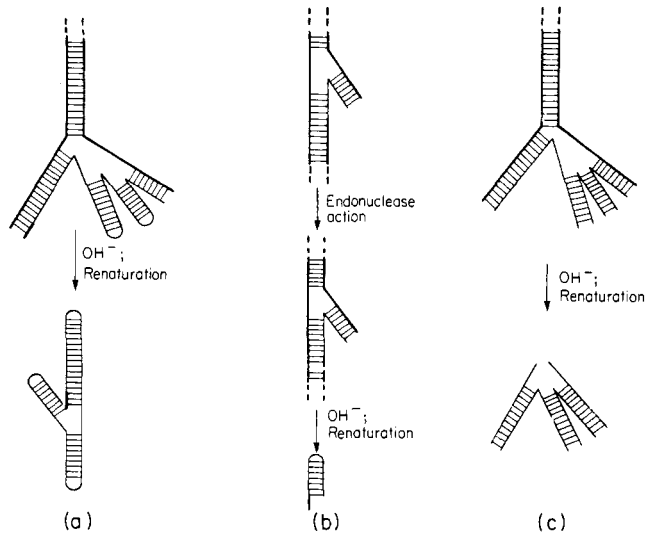

FIGURE 2: Three modes of generation of branched structures in the partial replication of duplex nucleic acids. Heavy lines represent the original template and lighter lines the nascent chains: (a) initiation of chains by an oligomeric primer; (b) initiation at a hydrolytic scission; (c) de novo chain synthesis. *No* significance is to be attached to the differences in the extent of branching depicted in the final separated products in the lower figures.

Included amongst these alternatives is Figure 2c with a mode of replication which entails initiation of new strands on the original template and, subsequently, upon those nascent strands as well. The replicated product (lower entry) is branched but multistranded, in contrast to the previously described models. The facile renaturability of this product would derive from the existence of multiple copies of the replicated termini of the double-strand template. In the current parlance, the low complexity of the product polynucleotide would be responsible for the rapid reassociation of these strands (Britten et al., 1974).

A more quantitive examination of the nature of the newly synthesized strands is worthwhile. For a 2.5-fold "replication", an electron micrographically determined number times length vs. length histogram was reported (Inman et al., 1965). A maximum in this distribution appeared to occur at a length of 1 μm—the length of the native T7 DNA was taken to be 11 μm.

Now, the most probable distribution represented by eq 12a has a maximum at $i^* \approx \langle i \rangle$ with $W^*(i) = (\langle i \rangle \mathrm{e})^{-1}$. If we refer the weight fraction of all species to $W(\langle i \rangle) = W^*(i)$, the distribution can be, in turn, written as

$$y = x \exp(1 - x) \qquad (16a)$$

where $y \equiv W(i)/W(\langle i \rangle)$ and $x \equiv i/\langle i \rangle$ are simple reduced variables. This distribution has constituted a reasonable trial description for many polymerization reactions (Stockmayer, 1974).

In Figure 3, the data referred to above is displayed in a semilogarithmic plot suggested by eq 16a. The points are experimental results and the straight line of unit slope is given by $-\ln(y/x) = x - 1$. These limited data describe a broad size range for the synthetic product. The inset of this figure contains a representation of the very narrow Poisson distribution in the following approximate analytical form

$$y = x^{-1/2} \exp[-\langle i \rangle (x - x \ln x - 1)] \qquad (16b)$$

For comparison with eq 16a, there is a semilogarithmic plot over the same range in the reduced variable $x - 1$. The pre-

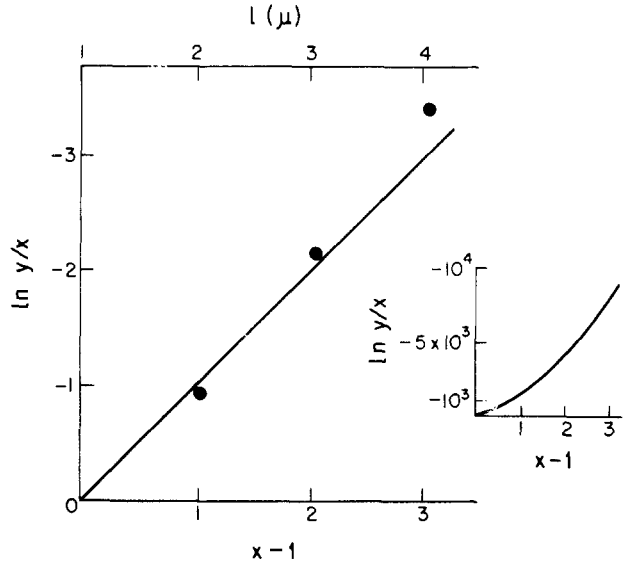

FIGURE 3: Plot of size population of the products of a partial replication of T7 DNA in vitro with $y \equiv W(i)/W^*(i) = W(i)/W(\langle i \rangle)$ and $x \equiv i/\langle i \rangle$ (see Figure 1). The upper scale records the total length of branched products in microns. Straight line conforms to the equilibrium most probable distribution. Data points are from Inman et al., 1965. Inset: A Poisson distribution over the same range of the reduced variable (x) and with a choice of $\langle i \rangle = 3 \times 10^3$ base pairs, corresponding to a duplex product of 1 μm length or 2×10^6 molecular weight.

cipitous decline of y from its relative maximum of unity is reflected in the large *negative* values of the logarithmic ordinate corresponding to this variation of the abscissa. Plainly, the broad population of chain lengths represented by the points in Figure 3 is not consistent with a Poisson distribution. Polynucleotide formation issuing from a fixed number of sites in Figure 2a,b might argue for a Poisson representation if depolymerization could be ignored. However, an increasing number of chain starts, as in Figure 2c, would contribute strongly toward a broadened spectrum of sizes even without a significant reverse reaction.

Expansion of the exponent in eq 16b about $x = 1$ and retention of terms in the second degree converts the Poisson distribution to one of Gaussian form (Feller, 1957).

$$y \approx \exp - [(\langle i \rangle/2)(x-1)^2] \qquad (16c)$$

An electron micrographic survey of naturally occurring DNAs up to the size of bacteriophage λ (inclusive of nucleic acids of the molecular weight of T7) indicates an apparently Gaussian fluctuation of the length about the mean size (Davis et al., 1971). With these lengths measured in units of the DNA of the small bacteriophage φX174 in its replicative or double-strand form, the standard deviation of the normal distribution for each specimen was found to vary with the square root of the mean length for that species with a common proportionality constant of 3.7×10^{-2}. The limit Gaussian of eq 16c has a standard deviation also proportional to the square root of the size but with a constant of $\langle i \rangle_\phi^{-1/2} = 1.4 \times 10^{-2}$, where $\langle i \rangle_\phi = 5.4 \times 10^3$ is the degree of polymerization in base pairs of the reference DNA of φX174 (Goulian et al., 1967). Thus, the DNAs of natural origin are more disperse as to size by a factor of almost three than would be expected from a Poisson process of replicative chain synthesis. Arguments can be constructed for duplication of a nucleic acid to experience size restrictions in addition to those applying to a relatively irreversible in vivo polymerization process. Consequently, this broader size spectrum may reflect unavoidable inimical cellular activity, e.g., nuclease action.

That the array of chain sizes of newly synthesized DNA replicated on a T7 nucleic acid template can be represented by a most probable distribution is, in very large measure, fortuitous. In part, this must reflect the mode of cataloging the branched product chains by their total length (Inman et al., 1965). Moreover, if a true equilibrium existed, the original template DNA and its pyrophosphorolyzed products should also be members of this broad population. The least controversial conclusion is that the range of lengths of the replicated chains is qualitatively very disperse and peaked at a size less than 10% of that of the original template.

These results invite comparison with a study of the in vitro replication of this bacteriophage DNA undertaken more recently (Hinkle and Richardson, 1974). This investigation utilized a system containing a DNA polymerase coded for by the T7 itself and augmented by a second gene product of this bacteriophage of then unknown function. Polymerization occurred without the addition of ribonucleoside triphosphates and, hence, apparently without the initiation of chains by an RNA polymerase activity. Most significantly, the synthetic product was not covalently attached to the template.

An electron microscopic analysis of the partially replicated template was appended (Bick, 1974). This companion study reported a number of characteristics of the product. The template associated with the newly synthesized chains exhibited "forks" though none are evident in the original T7 nucleic acid. The duplex arms of these forks grew to a size of only about 2×10^6 daltons and the size distribution became broader with increasing time.

These two investigations almost a decade apart led to a partially replicated T7 DNA of approximately the same size and size range with some similar properties if forks are seen as equivalent to branches. There is an abundance of reports of such partial replication of DNA from bacterial viruses of intermediate and large sizes where the experimental systems employed polymerases of different origin and gene products of the virus or cell deemed to be required for replication. However, a different picture emerges when *cell extracts* from *E. coli* provide the media for template replication. For example, synthetic products with molecular weights of that of T7 DNA are formed (Masker and Richardson, 1976a) which are biologically active, as judged by susceptibility of spheroplasts to infection (Masker and Richardson, 1976b). The proposition developed here is that the unsupplemented systems are essentially deficient thermodynamically. To rectify this lack, the pyrophosphate by-product of polynucleotide synthesis must be hydrolytically removed, as most certainly happens in the cell and very probably in the cell extract systems as well.

Conclusion

The experimental results discussed in the previous section were interpreted as representing an incomplete replication which was attributed to the onset of pyrophosphorolysis in turn capable of elimination by the action of inorganic pyrophosphatase. That in vitro nucleic acid synthesis should be joined to pyrophosphate hydrolysis as it is in vivo is a surprisingly heterodox view. Most biochemistry and biology texts point to the advantages, if not the necessity, of this process in cellular metabolism. In the specific area of nucleic acid synthesis, a widely read introductory survey of molecular biology notes with fulgent regularity that for reaction 2 above "ΔG is slightly positive (~0.5 kcal/mol). This immediately poses the question, since polynucleotides obviously form: What is the source of the

free energy? The needed free energy arises from the splitting of the high-energy pyrophosphate group which is formed simultaneously with the high-energy phosphodiester bond" (Watson, 1965, 1970, 1976).

Now, the free-energy change for reaction 2, as stated above, is negative but with a magnitude strongly conditional on base-stacking interactions. Moreover, polynucleotides of considerable size do form in vitro *without* any accompanying hydrolysis of pyrophosphate. Three possible explanations for this behavior have been presented here. Nonetheless, the conclusion of the above quotation does seem sustainable if, in addition to the influence of this coupled reaction on the ceiling for product size, its effect on the persistence of a narrow size distribution is also considered.

References

Alberty, R. A. (1958), *J. Am. Chem. Soc. 80,* 1777.

Alberty, R. A. (1969), *J. Biol. Chem. 244,* 3290.

Applequist, J., and Damle, V. (1965), *J. Am. Chem. Soc. 87,* 1450.

Beyersmann, D., and Schramm, G. (1968), *Biochim. Biophys. Acta 159,* 64.

Bick, M. D. (1974), *J. Biol. Chem. 249* (appendix to Hinkle, and Richardson).

Bloomfield, V., Peller, L., and Alberty, R. A. (1962), *J. Am. Chem. Soc. 84,* 4367.

Bollum, F. (1962), *J. Biol. Chem. 237,* 1945.

Bollum, F. (1974), *Enzymes, 3rd Ed., 10,* 157–160.

Borer, P. N., Dengler, B., Tinoco, I. Jr., and Uhlenbeck, O. (1974), *J. Mol. Biol. 86,* 843.

Bremer, H. (1967), *Mol. Gen. Genet. 99,* 362.

Bremer, H., Konrad, M. W., Gaines, K., and Stent, G. S. (1965), *J. Mol. Biol. 13,* 540.

Britten, R. J., Graham, D. E., and Neufeld, B. R. (1974), *Methods Enzymol. 29,* 363.

Brown, W. B., and Szwarc, M. (1958), *Trans. Faraday Soc. 54,* 416.

Chamberlin, M. (1974), *Enzymes, 3rd Ed. 10,* 353–366.

Davis, R. W., Simon, M., and Davidson, N. (1971), *Methods Enzymol. 21D,* 424.

Deutscher, M. P., and Kornberg, A. (1969), *J. Biol. Chem. 244,* 3019.

Dunn, J. J., and Studier, F. W. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 1559.

Feller, W. (1957), An Introduction to Probability Theory and Its Applications, 2nd ed, Vol. 1, New York, N.Y., Wiley, p. 176.

Flory, P. J. (1940), *J. Am. Chem. Soc. 62,* 1561.

Flory, P. J. (1953), Principles of Polymer Chemistry, Ithaca, N.Y., Cornell University Press, chapter VIII.

Geiduschek, E. P., and Haselkorn, R. (1969), *Annu. Rev. Biochem. 38,* 652.

Godefroy-Colburn, T., and Grunberg-Manago, M. (1972), *Enzymes, 3rd Ed. 7,* 566.

Gold, L. (1958), *J. Chem. Phys. 28,* 91.

Goulian, M., Kornberg, A., and Sinsheimer, R. L. (1967), *Proc. Natl. Acad. Sci. U.S.A. 58,* 2321.

Grunberg-Manago, M. (1963), *Prog. Biophys. Mol. Biol. 13,* 191.

Grunberg-Manago, M., and Ochoa, S. (1955), *J. Am. Chem. Soc. 77,* 3165.

Hayes, F. N., Mitchell, V. E., Ratliff, R. L., and Williams, D. L. (1967), *Biochemistry 6,* 2488.

Hinkle, D. C., and Richardson, C. C. (1974), *J. Biol. Chem. 249,* 2974.

Hyman, R. W., and Davidson, N. (1970), *J. Mol. Biol. 50,* 421.

Inman, R. B., Schildkraut, C. L., and Kornberg, A. (1965), *J. Mol. Biol. 11,* 285.

Kato, K-I., Gonçalves, J. M., Houts, G. E., and Bollum, F. J. (1967), *J. Biol. Chem. 242,* 2780.

Kavenoff, R., and Zimm, B. H. (1973), *Chromosoma 41,* 1.

King, E. L., and Altman, C. (1956), *J. Phys. Chem. 60,* 1375.

Kornberg, A. (1962), in Horizons in Biochemistry, Kasha, M., and Pullman, B., Ed., New York, N.Y., Academic Press, p. 251.

Kornberg, A. (1965), in Evolving Genes and Proteins, Bryson, V., and Vogel, J., Ed., New York, N.Y., Academic Press, pp 403–417.

Kornberg, A. (1974a), DNA Synthesis, San Francisco, Calif., W. H. Freeman, p 41.

Kornberg, A. (1974b), DNA Synthesis, San Francisco, Calif., W. H. Freeman, p 111.

Kornberg, A., Bertsch, L. L., Jackson, J. F., and Khorana, H. G. (1964), *Proc. Natl. Acad. Sci. U.S.A. 51,* 315.

Krakow, J. S., Coutsogeorgopoulos, C., and Cannellakis, E. S. (1961), *Biochem. Biophys. Res. Commun. 5,* 477.

Krakow, J. S., Daley, K., and Fronk, E. (1968), *Biochem. Biophys. Res. Commun. 32,* 98.

Masker, W. E., and Richardson, C. C. (1976a), *J. Mol. Biol. 100,* 543.

Masker, W. E., and Richardson, C. C. (1976b), *J. Mol. Biol. 100,* 557.

Miyake, A., and Stockmayer, W. H. (1965), *Makromol. Chem. 88,* 90.

Morawetz, H. (1965), Macromolecules in Solution, New York, N.Y., Interscience, p 13.

Nanda, V. S., and Jain, S. C. (1970), *Eur. Polym. J. 6,* 1517.

Oosawa, F. (1970) *J. Theor. Biol. 27,* 69.

Oosawa, F., and Kasai, M. (1962), *J. Mol. Biol. 4,* 10.

Peller, L. (1966), *Proc. Natl. Acad. Sci. U.S.A. 55,* 1025.

Peller, L. (1975), *Biochem. Biophys. Res. Commun. 63,* 912.

Peller, L. (1976a), *J. Phys. Chem. 80,* 2462.

Peller, L. (1976b), *Biochemistry 15,* 141.

Peller, L., and Barnett, L. (1962), *J. Phys. Chem. 66,* 680.

Radding, C. M., Josse, J., and Kornberg, A. (1962), *J. Biol. Chem. 237,* 2869.

Schekman, R., Weiner, A., and Kornberg, A. (1974), *Science 186,* 987.

Sela, M., and Katchalski, E. (1958), *Adv. Protein Chem. 13,* 302–307.

Smith, D. A., Ratliff, R. L., Williams, D. L., and Martinez, A. M. (1967), *J. Biol. Chem. 242,* 590.

Solie, T. N., and Schellman, J. A. (1968), *J. Mol. Biol. 33,* 61.

Spiegelman, S., Haruna, I., Holland, I. B., Beaudreau, G., and Mills, D. (1965), *Proc. Natl. Acad. Sci. U.S.A. 54,* 919.

Stockmayer, W. H. (1974), *Science 186,* 724.

Tanford, C. (1961), Physical Chemistry of Macromolecules, New York, N.Y., Wiley, p 605.

Watson, J. D. (1965), Molecular Biology of the Gene, 1st ed, New York, N.Y., W. A. Benjamin, pp 155–156.

Watson, J. D. (1970), Molecular Biology of the Gene, 2nd ed, New York, N.Y., W. A. Benjamin, pp 155–156.

Watson, J. D. (1976), Molecular Biology of the Gene, 3rd ed, Menlo Park, Calif., W. A. Benjamin, p 125.

Zimm, B. H. (1960), *J. Chem. Phys. 33,* 1349.